

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 11:08:20 ; Search time 13.29 Seconds  
(without alignments)  
1467.697 Million cell updates/sec

Title: US-09-911-513-2

Perfect score: 2758

Sequence: 1 MKRDHHHHQDKKTKMMNEE.....MLGWHTRPLIATSAWKLSTN 532

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	112.5	4.1	772	1 NFL1_HUMAN	Q14494 h nuclear f
2	107	3.9	1158	1 CNDL_SCHPO	Q34679 schizosacch
3	106.5	3.9	1420	1 YMBB_YEAST	Q03496 saccharomyc
4	106	3.8	3206	1 POLG_PSBMV	P29152 p genome po
5	104.5	3.8	953	1 LKAB_PASHA	P55116 pasteurella
6	104.5	3.8	4568	1 DYRC_CAEEL	Q19020 caenorhabdi
7	102.5	3.7	953	1 LKAB_PASHA	P55118 pasteurella
8	101.5	3.7	580	1 RUBA_CHLRE	Q42694 chlamydomon
9	101.5	3.7	953	1 LKAB_PASHA	P16535 pasteurella
10	101	3.7	414	1 SPYA_CALJA	P10299 callithrix
11	100	3.6	1007	1 MA2B_FELCA	Q46432 felis silve
12	99.5	3.6	1181	1 YY02_METJA	Q60301 methanococc
13	99.5	3.6	1436	1 DPO3_STAAU	Q53665 staphylococ
14	99	3.6	414	1 SPYA_FELCA	P41689 felis silve
15	99	3.6	636	1 PHBC_RHET	Q52728 r poly-beta
16	98.5	3.6	558	1 ORC2_XENLA	Q19628 xenopus lae
17	98.5	3.6	4590	1 FATH_HUMAN	Q14517 homo sapien
18	97.5	3.5	1013	1 A60D_DROME	P91927 drosophila
19	96.5	3.5	761	1 PQOF_KLEPN	P27508 klebsiella
20	96.5	3.5	851	1 OBP_HSV11	P10193 herpes simp
21	96	3.5	350	1 SUCA_PIG	Q19069 sus scrofa
22	96	3.5	546	1 PGMO_ECOLI	P36938 escherichia
23	96	3.5	787	1 Y091_CAEEL	P41842 caenorhabdi
24	95	3.4	527	1 TCPB_SCHPO	Q10147 schizosacch
25	95	3.4	665	1 ATKB_THEAC	P57700 thermoplas
26	95	3.4	1549	1 GLSF_CYACA	Q19906 cyanidion c
27	94.5	3.4	614	1 Y156_ARCFU	Q28422 archaeoglob
28	94.5	3.4	1260	1 YAOE_SCHPO	Q10093 schizosacch
29	94.5	3.4	1597	1 CTRO_MOUSE	P49025 mus musculu
30	94	3.4	1071	1 CARB_BACSU	P25994 bacillus su
31	94	3.4	4563	1 APB_HUMAN	P04114 homo sapien
32	93.5	3.4	877	1 DPO1_STRPN	P13252 streptococc
33	93.5	3.4	1330	1 VCAP_PRVIS	Q00705 pseudorabie

RESULT 1

ID	NFL1_HUMAN	STANDARD;	PRT;	772 AA.
AC	Q14494; 012877;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	NUCLEAR FACTOR ERYTHROID 2 RELATED FACTOR 1 (NF-E2 RELATED FACTOR 1)			
DE	(NF-E2-RELATED FACTOR 1) (NUCLEAR FACTOR, ERYTHROID DERIVED 2, LIKE 1)			
DE	(TRANSCRIPTION FACTOR 11) (TRANSCRIPTION FACTOR HBZ17) (TRANSCRIPTION			
DE	FACTOR LCR-F1) (LOCUS CONTROL REGION-FACTOR 1).			
GN	NFE2L1 OR NRE1 OR TCF11 OR HBZ17.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=95095252; PubMed=8001966;			
RA	Luna L., Johnsen O., Skartlien A.H., Pedetour F., Turc-Carel C.,			
RA	Prydz H., Kolstoe A.-B.;			
RT	"Molecular cloning of a putative novel human bZIP transcription			
RT	factor on chromosome 17q22."			
RL	Genomics 22:553-562(1994).			
RL	[2]			
RP	SEQUENCE OF 326-772 FROM N.A.			
RP	MEDLINE=94310069; PubMed=8036168;			
RA	Caterina J.J., Donze D., Sun C.W., Ciavatta D.J., Townes T.M.;			
RT	"Cloning and functional characterization of LCR-F1: a bZIP			
RT	transcription factor that activates erythroid-specific, human globin			
RT	gene expression."			
RL	Nucleic Acids Res. 22:2383-2391(1994).			
CC	-!- FUNCTION: ACTIVATES ERYTHROID-SPECIFIC, GLOBIN GENE EXPRESSION.			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed, usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X77366; CAA54555.1; -			
DR	EMBL; U08853; AAA20466.1; -			
DR	HSP; P34707; LSKN.			
DR	MIM; 163260; -			
DR	InterPro; IPR001871; bZIP.			
DR	Pfam; PF00170; bZIP; 1.			
DR	SMART; SM00338; BRZL; 1.			
DR	PROSITE; PS00036; bZIP_BASIC; 1.			
KW	Transcription regulation; Activator; DNA-binding; Nuclear protein.			
FT	DOMAIN 125 288 ASP/GLU-RICH (ACIDIC).			
FT	DOMAIN 496 517 POLY-SER.			
FT	DNA_BIND 659 674 BASIC MOTIF.			

Q64535 rattus norv  
P31030 oryctolagus  
O04015 a delta l-p  
O84303 chlamydia t  
Q43933 homo sapien  
P07210 human rhino  
P26595 mus caroli  
P00561 escherichia  
P53254 saccharomyc  
P10935 rhizobium m  
Q29052 sus scrofa  
O13776 schizosacch





DR Pfam; PF00270; DEAD; 1.  
DR Pfam; PF00863; Peptidase\_C4; 1.  
DR Pfam; PF00851; Peptidase\_C6; 1.  
DR Pfam; PF00767; Poty\_coat; 1.  
DR Pfam; PF01577; Poty\_P1; 1.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR PRINTS; PR00966; NIAPOTYPITASE.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELIC; 1.  
KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;  
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;  
KW ATP-binding.  
FT CHAIN 1 ? N-TERMINAL PROTEIN.  
FT CHAIN ? 856 HELPER COMPONENT PROTEINASE.  
FT CHAIN 857 ? PROTEIN P3.  
FT CHAIN ? 1266 6 KDA PROTEIN 1.  
FT CHAIN 1267 1902 CYTOPLASMIC INCLUSION PROTEIN.  
FT CHAIN 1903 1955 6 KDA PROTEIN 2.  
FT CHAIN 1956 ? GENOME-LINKED PROTEIN.  
FT CHAIN ? 2395 NUCLEAR INCLUSION PROTEIN A.  
FT CHAIN 2396 2915 NUCLEAR INCLUSION PROTEIN B.  
FT CHAIN 2916 3206 COAT PROTEIN.  
FT SITE 1266 1267 CLEAVAGE (BY 49 KDA PROTEASE).  
FT SITE 1902 1903 CLEAVAGE (BY 49 KDA PROTEASE).  
FT SITE 1955 1956 CLEAVAGE (BY 49 KDA PROTEASE).  
FT SITE 2395 2396 CLEAVAGE (BY 49 KDA PROTEASE).  
FT SITE 2915 2916 CLEAVAGE (BY 49 KDA PROTEASE).  
FT BINDING 2016 2016 COVALENT LINKAGE OF VIRAL RNA (BY SIMILARITY).  
FT NP\_BIND 1351 1358 ATP (POTENTIAL).  
FT SEQUENCE 3206 AA; 364271 MW; 42A3D921B5A0CBF CRC64;  
Query Match 3.8%; Score 106; DB 1; Length 3206;  
Best Local Similarity 20.08; Pred. No. 40;  
Matches 120; Conservative 84; Mismatches 170; Indels 226; Gaps 33;  
Qy 8 HQDKKMM--NEEDGNGM-DELLAVLYK-----VRSEMAV 45  
Db 70 HLVDKTPKRGSOFTGDMAAEIKVTGAEPVNCFMVGTICKINENSVKGVMAAI 129  
Qy 46 AOKLEQVMSNVQEDDLSLATETVHNPAAE---LYTWLDSMLTDLNPPSSNAEYDK 102  
Db 130 PRQITQDEVFMRKARLQ--AAVASTIEREKEKQFAESKLEELRARREKLKG-IVIK 186  
Qy 103 AIPG-----DAILNQAIDSASSNGGGGDT-----YTTNKRKCSNGVVETTTA---- 148  
Db 187 TRKGLEWREATPNQGRKLQSTFSDASGGKTLTPHTIYCKTKSKFSNGVKCATSKMR 246  
Qy 149 -----TAESTRHVIVDSQE-----NGVRLV----- 169  
Db 247 TVRKPSLKMKTESIDVLIEQVMTIAGKHAKOVLTDKQKTRNVIRVNGVRLQLQVETK 306  
Qy 170 -HALLACAEAVQKENVLVAEALVKQIGFVLSQVIGAMRKVATYFAEALRIYRLSPQS 228  
Db 307 HKGIISQKDALNNLT-----KRVARHF--ARKTAIHPSDS 342  
Qy 229 PIDHSLDSTLQMHYETCPYILKFAHFTANQAI-----LEAFQK-----KRV 270  
Db 343 -ITHGSGV-----FLR-ANISGSKSYSIDDLFFVVRGKRNGKLMSRNKVAWRKM 391  
Qy 271 HVIDFSMSOGLQ-WPAL-MOALALRP-----GG-----PPVPERL 302  
Db 392 FOIDHFSIVGIRKINWAFDAEYVVKURDESVDSDHDCVGGITPECGILAAQILRVFPCWRI 451  
Qy 303 T-----GIGPPAPDNFDNYLHVEYGCKLA-----HLAEAIHVEFEYRGFVANTLADLA 349  
Db 452 TCTKCISNWLKSKPTSEQIEHIIYERG-NLAIQDLNKRIPSAHV-----T 494  
Qy 350 SMLELRSEETESAVNSVFEL-----HKLGRPGCATDKVLGVNVQIKPETF-----TVV 398  
Db 495 QMVELLRQRIK-----NTTFDMGNNTKVHELIGH--RODGVFRHLNRLNNSILAANGSSTI 548

Qy 399 EQESHNSPIFLDRP-----TESLHYSTLFDLSLEGVPSQDKVMSEVYLKQICNVVACD 454  
Db 549 EWESNMSELLELARWHNKRRTESI-----ASGGISSFRNKRISAKA-----QINFALMCD 596  
RESULT 5  
LKA3\_PASHA  
ID LKA3\_PASHA STANDARD; PRT; 953 AA.  
AC P55116;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE LEUKOTOXIN FROM SEROTYPE T3.  
GN LKTA.  
OS Pasteurella haemolytica.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Mannheimia.  
OX NCBI\_TaxID=75985;  
(1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=SEROTYPE T3;  
RC MEDLINE=94041617; PubMed=8225575;  
RA Burrows L.L., Lo R.V.C., Olah-Winfield E.;  
RT "Molecular analysis of the leukotoxin determinants from Pasteurella haemolytica serotypes 1 to 16."  
CC Infect. Immun. 61:5001-5007(1993).  
CC -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY DEFINED.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC ACTIVITY.  
CC -!- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).  
CC -!- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN MODIFIED (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U01216; AAB36691.1; -.  
DR HSSP; P02392; ICTF.  
DR InterPro; IPR001343; Hemlysn\_Ca\_bind.  
DR InterPro; IPR003355; RTX\_N.  
DR Pfam; PF00353; hemolysinCabin; 1.  
DR Pfam; PF02382; RTX; 1.  
DR PROSITE; PS00330; HEMOLYSIN\_CALCIIUM; 2.  
DR Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium; Transmembrane; Lipoprotein; Palmitate.  
KW TRANSMEM 229 249 POTENTIAL.  
FT TRANSMEM 297 318 POTENTIAL.  
FT TRANSMEM 381 401 POTENTIAL.  
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH.  
FT REPEAT 734 739 1.  
FT REPEAT 743 748 2.  
FT REPEAT 752 757 3.  
FT REPEAT 761 766 4.  
FT REPEAT 770 775 5.  
FT REPEAT 779 784 6.  
SQ SEQUENCE 953 AA; 101948 MW; FDBDCE2FDC85FDF2 CRC64;  
Query Match 3.8%; Score 104.5; DB 1; Length 953;  
Best Local Similarity 18.8%; Pred. No. 8.8;  
Matches 124; Conservative 97; Mismatches 199; Indels 239; Gaps 33;

Qy 19 BEDDNGMDLAV---LGYKVRSEMAVDAVK-----LEQLEYVMNVQEDD 63  
Db 55 DSGSGNGQLDLVKAEEELGIEVQKEEGNDIAKAQTSCTIQNLGLTGERGIVLAPOLDK 114  
Qy 64 LSQI---ATETVHYNPAELYTWLDSMLTDLNPPSSNAEYDLKAIPGDAILNQFAI 115  
Db 115 LQKNKVGQALGSSSEIAQNSQAKTVLSGVOSILGSVLGMDLD---EALONE--- 165  
Qy 116 DSASSNSGGGGDYVTYTKR-LKCSNGVETTT----- 147  
Db 166 -----SDQLTLAKAGLELTNSLIENANSVQTLDAFSEQISQFGSKLQNVKGLG 214  
Qy 148 ATASTRHVLVDSOENGVRVLVHALLACAEAVQENITVAE---ALVKQIGF---LAVSQ 201  
Db 215 ALGOKLKNIGGLDAGLHVISGLLSGATAA---LVLDKADASTAKKVGAGFELANQV 270  
Qy 202 IGAMRK-VATYFAEALABRIYRLSPSPIDHLSLSDTLQMHFYETCPYLKFAHATANQAI 260  
Db 271 VGNITKAVSSVI---LAORVAARLSSTGTPVAALIASTVAL---AISPLSFA-----GI 317  
Qy 261 LEAFQGGKRVH-----VIDFSMSQGLQWPMALMOALALRPG 295  
Db 318 ADKFDRAKSLNAYAEFKKLYEGDSLAEYOHGTGTIDASVT-----AINTALAAIAG 371  
Qy 296 G-----PPVFLRTGIGPPAPDNFDYLVHEVGCKLAHLAEAIH---VEFE----- 335  
Db 372 GVSAAAGSVVASPTALVSGITGVISTILQYSKQ---AMFEHVANKIHNKIVEKNNGG 429  
Qy 336 ---YRGVANTLADLASM---LELRPSEIESVAVNSVFE-----LHKLGRPGALDK 382  
Db 430 KNYPENGYDARYLANLQNMKFLNLN-KELQAEVIAITOOQWDSNIGDLGAGISRLGEK 488  
Qy 383 VLGVVNOIKPEIFTVVEGESNHNSPIFLDRFTESLHYYS---TLFDSLEGV---PSQD 435  
Db 489 VL-----SGKAYVDAFEQGHKLKADKLVLQDSAKGIIDVSNTEA 528  
Qy 436 KVMSEVYLGKQICNVVACDGPDRVERHET-----LSQWRNRFGSA----- 475  
Db 529 KTOHILF-----TPLLTPGTGKERVQTGKYEYITKLHINRVDSWQTKDGAASSTDLT 583  
Qy 476 -----GFAAAHIGSNAP--KOASMLLAL-----FNGEGY-RVEESDG 510  
Db 584 NVQIRIGVELDH-AENVIKTKETKIVATLGDGDDNVFVSGSTTEIDGEGYDRVHYSRG 641  
RESULT 6  
ID DYHC CAEL STANDARD; PRT; 4568 AA.  
AC Q19020;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE DYNEIN HEAVY CHAIN, CYTOSOLIC (DYHC).  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoida;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN 1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=96114101; PubMed=8674131;  
RA Lye R.J., Wilson R.K., Waterston R.H.;  
RT "Genomic structure of a cytoplasmic dynein heavy chain gene from the  
RL nematode Caenorhabditis elegans";  
RL Cell Motil. Cytoskeleton 32:26-36(1995).  
CC -!- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A  
CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND  
CC ORGANELLES ALONG MICROTUBULES.  
CC -!- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF  
CC INTERMEDIATE AND LIGHT CHAINS.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; L33260; AAC37251.1; -.  
DR InterPro; IPR003593; AAA.  
DR SMART; SM00382; AAA; 1.  
KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.  
FT DOMAIN 587 652 COILED COIL (POTENTIAL).  
FT DOMAIN 814 844 COILED COIL (POTENTIAL).  
FT DOMAIN 1241 1274 COILED COIL (POTENTIAL).  
FT DOMAIN 1324 1340 COILED COIL (POTENTIAL).  
FT DOMAIN 1559 1591 COILED COIL (POTENTIAL).  
FT DOMAIN 1964 1992 MICROTUBULE-BINDING (POTENTIAL).  
FT DOMAIN 3132 3229 COILED COIL (POTENTIAL).  
FT DOMAIN 3339 3432 COILED COIL (POTENTIAL).  
FT DOMAIN 3707 3739 COILED COIL (POTENTIAL).  
FT DOMAIN 4359 4386 COILED COIL (POTENTIAL).  
FT NP\_BIND 1865 1872 ATP (POTENTIAL).  
FT NP\_BIND 2163 2170 ATP (POTENTIAL).  
FT NP\_BIND 2537 2544 ATP (POTENTIAL).  
FT NP\_BIND 2880 2887 ATP (POTENTIAL).  
SQ SEQUENCE 4568 AA; 521568 MW; 028E52684F381676 CRC64;

Query Match 3.8%; Score 104.5; DB 1; Length 4568;  
Best Local Similarity 22.2%; Pred. No. 87;  
Matches 105; Conservative 47; Mismatches 145; Indels 177; Gaps 24;  
Qy 113 FAIDSSASSNSGGGGDYTTNKRLLKCSNGVYETTTATAESTRHVLVDSOENGVRVH-- 170  
Db 3644 FTVTSSLSAQ-----CLNQVLRSEPPVDKRRNDLLKQGEFAVRLHLE 3689  
Qy 171 -ALLACAEAVQKEML---TVAAELVQKIGFLA-VSQIGAMRKVATYFAEALARIYRLSP 225  
Db 3690 KALLAALNESKGLTDDNSVETLEKLNKAEAAVQAKSAETDKVMAEVDVAVSAQYQRLST 3749  
Qy 226 SQSPIDHLSDTLQMHF--YETCPYLK--FAHFTANQAILEAFQGGKRVHVIDFSMSQGL 281  
Db 3750 ACSHIYHTLQOLNEIHFYHSLDFEIVFTHVLKTELSSTTDYAKRLRIITSLFQ-- 3807  
Qy 282 QWPMALQALALRPGGPPVFLRTGIGPPAPDNFDYLVHEVGCKLAHLAEAIHVEFEYRGFVA 341  
Db 3808 -----TVFRVSRG-----MLHTDKVLLALLMLRIHR----- 3835  
Qy 342 NTLADLDSMLRLPSEIESVAVNSVFEHLKLLGR-----PGAID----- 381  
Db 3836 -----SNPSAPAYEQHFDL--LLGRSDFVAKNDEADSTIPGGLDFTVEN 3878  
Qy 382 -----KVLGVVNOIKPEIFTVVEGES-----NHN-----SPIF 409  
Db 3879 KKSIAKARKVVGFEF-----VFAHLQNSAAVTSWLTNDNPESNPVYVWDADGKLSPLC 3933  
Qy 410 L-----DRETESLH-VYSTLFDS--LEGVPSGQDKVM-----SEV 441  
Db 3934 IAMNSLLVHVALRPDLRLMASAHRVVSTAFDDHFMQ-----QDKVVDILSIVDNEVSPSEP 3988  
Qy 442 YLKGQICNVVACDGPDRVERHETLSQWRNRFGSAGFAAHIGS-NAFKQASMLL 494  
Db 3989 VL---LCSATGYDASGKI---EDLAVETNR-----QLTSLAIGSAEGFNQADSAL 4032  
RESULT 7  
LKAB\_PASHA  
ID LKAB\_PASHA STANDARD; PRT; 953 AA.  
AC P55118;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)



```
DR Pfam: PF00118; cpn60_TCPI1; 1.
DR PRINTS: PR00298; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00296; CHAPERONINS.CPN60; 1.
KW Chapterone; ATP-binding; Chloroplast; Transit peptide; Heat shock.
FT TRANSIT 1 580 CHLOROPLAST (POTENTIAL).
FT CHAIN ? 580 RUBISCO SUBUNIT BINDING-PROTEIN ALPHA
FT SUBUNIT.
SQ SEQUENCE 580 AA; 61863 MW; 16FD34B115E7067F CRC64;

Query Match 3.7%; Score 101.5; DB 1; Length 580;
Best Local Similarity 21.7%; Pred. No. 7.1;
Matches 100; Conservative 45; Mismatches 129; Indels 187; Gaps 24;

QY 65 SOLATETVHNPALYTWLDSMLTDLNPPS-----SNAEY-----DLKAIPG-DAIL 110
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 126 SVLAREMIHYG-----LQSVTAGANPIAVKRGDKTAEVLVAKLKEHAPVKGRRDIK 178
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 111 NOFAIDSASSNOGGGGDTYTNKRLKCSNGV--VETTTAT-----
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 179 N---VASISAGNDNAIGEMADALDKVGSNGVLSTSTSTVVEVQEGMEIDRGYISP 235
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 150 -----AESTRHVVLVDSQ-----ENGVRVLVHALLACAEAVQKLENL--- 184
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 236 QFVTNQERLLVEYDNCRLVTDQKIDAIRDIPILEQVTRNLNAPLLIIAEDVSGEALATL 295
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 185 -----TVAEALVKOIGF-----LAVSOIGAMRK 207
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 296 VVNLKRGVLNVCAKAPGFGERRKSLLODIAIVTGAEFIAKDGMKVQDAVVEQLGVARK 355
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 208 V-----ATYFAEALAR-----RIYRLSPQSPIDHSLSDTLQMHFYETCPYLFAHPTA 256
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 356 VTVANNTTLLIADAASKDEIEMRIAQLKELAETD-SYVDTEKLS-----ERIAKLSG 407
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 257 NQAIL-----EAFGKKRVHVHVFDSMSQGLQWLPALMQALALRPGGPPVFLTGIGPPAP 310
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 408 GVAVIKVGAAYTAELEDRKRIED---AKNATFAAVEGIV--PGG-----
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 311 DNFDYLHEVGCKLAHLAEAIHVEFYRGFVANTLAD-----LDASMLELR-PSEI--ES 361
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 449 -----GAALLHLSLVPAFRE-----TLTDAEEKLGADIVKMSLRAPCRLIADN 492
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 362 VAVNSVFELHKLGRP-----GAIDKV-----LGVVNQIK 391
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 493 AGVEGEIVQRLGKPFEGVGYNAMIDKVENLLDAGVIDPAK 533
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 9
ID LKAL_PASHA STANDARD; PRT; 953 AA.
AC P16535;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 30-MAR-2000 (Rel. 39, Last annotation update)
DE LEUKOTOXIN FROM SEROTYPE A1.
GN LKTA.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Mannheimia.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE A1;
RX MEDLINE=87306837; PubMed=3040588;
RA Lo R.X.C., Strathdee C.A., Shewen P.E.;
RT "Nucleotide sequence of the leukotoxin genes of Pasteurella
   haemolytica A1."
RL Infect. Immun. 55:1987-1996(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE A1 / PHL101;
RX MEDLINE=89210283; PubMed=2707120;
```

```
RA Highlander S.K., Chidambaram M., Engler M.J., Weinstock G.M.;
RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene
   cluster."
RL DNA 8:15-28(1989).
RN [3]
RP SEQUENCE OF 884-953 FROM N.A.
RC STRAIN=SEROTYPE A1 / PHL101;
RX MEDLINE=90236888; PubMed=2185213;
RA Highlander S.K., Engler M.J., Weinstock G.M.;
RT "Secretion and expression of the Pasteurella haemolytica Leukotoxin."
RL J. Bacteriol. 172:2343-2350(1990).
CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
   CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
   DEFINED.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
   CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
   ACTIVITY.
CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
   INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -1- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN
   MODIFIED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (see http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC
DR EMBL: M20730; AAA25529.1; -
DR EMBL: M24197; AAA25543.1; -
DR PIR: S29516; S29516.
DR HSP; P02392; ICTF.
DR InterPro; IPR001343; Hemlysn_Ca_bind.
DR InterPro; IPR003355; RTX_N.
DR Pfam; PF00353; hemolysinCabin; 1.
DR Pfam; PF02382; RTX; 1.
DR PRINTS: PR00313; CABNONGRPT.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 4.
KW Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
   Transmembrane; Lipoprotein; Palmitate.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 297 318 POTENTIAL.
FT TRANSMEM 366 390 POTENTIAL.
FT DOMAIN 734 784 6 X REPEATS, GLY-RICH (BY SIMILARITY).
FT REPEAT 734 739 1.
FT REPEAT 743 748 2.
FT REPEAT 752 757 3.
FT REPEAT 761 766 4.
FT REPEAT 770 775 5.
FT REPEAT 779 784 6.
FT CONFLICT 409 414 FEHVN -> LSTLQI (IN REF. 2).
FT CONFLICT 742 742 D -> Y (IN REF. 2).
SQ SEQUENCE 953 AA; 101996 MW; 7F93D113A118C05F CRC64;

Query Match 3.7%; Score 101.5; DB 1; Length 953;
Best Local Similarity 18.8%; Pred. No. 15;
Matches 122; Conservative 100; Mismatches 207; Indels 221; Gaps 32;

QY 19 EEDDNGMGDELLAV---LGYKVRSSSEMAVQAOK-----LEQLEVMNSNVQEDD 63
   : : ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 55 DTGNGNQLDLVKAEEELGIEVQREERNNTATQTSIGTQTQTAIGLTERGIVLSAPQIDK 114
   : : ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 64 LSQ-----ATETVHNPALYTWLDSMLTDLNPPSSNAEYDLKAIPGDAILNOFAI 115
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 115 LLQKTKAGQALGSAESIVQNAKAKTVLSGIQSLGSLVAGMDLD-EALQNN--NQHAL 171
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 116 DSASSNQGSGGDTYTNKRLKCSNGVYVTTT-----A 148
   ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
```



```
Db 172 AKAG-----LELTNSLIENTANSVKTIDFEGEQISQSGKLNQIKGLGT 215
Qy 149 TAESTRHVVLDQSGVRLVHALLACAEAVQKENLTVAE---ALVKQIGF---LAVSQI 202
Db 216 LGDKLNIGGLDKAGLDVLSGLSGATAA---LVLADKNASTAKVKVAGFELANQVV 271
Qy 203 GAMRK-VATYFAEALARIYRLSPSQSDISLSDTLOMHFYETCPYL-----KFAHPT 255
Db 272 GNITRAVSSYI---LAQRAAGLSSTGPVAALIASTVSL---AISPLAFAGTADKFNHAK 325
Qy 256 ANOATLEAF-----QGKKRVHVDFMSOGQLWPALMOALALRPGG-----296
Db 336 SLESTAEFEKKGCDGDLNLLAEYQGTGTIDASVI-----AINTALAAIAGGVSAAAAG 379
Qy 297 -----PPVFRITGIPGPDNFYDLYHEVGCKLAHLAEAIH---VEFE-----YRGF 339
Db 380 SVIASPIALLVSGITGVISTILQYSKQ--AMFEHVANKIHKNIVEWKNHGNKFYFENG 437
Qy 340 VANTLADLDASM---LELRPSEIESVAVNSVFE-----LHKLGRPCAIDKVLGVVNOI 390
Db 438 DARYLANLDQNMKFLNLN-KELQAEYVIAITQQQWDMNIGDLAISRLGKVL-----490
Qy 391 KPEIFTVVEQSNHNSPIFLDRFETSLHYVS---TLFDSLEGV---PSGODKYMSEVYL 443
Db 491 -----SGKAYVDFAFEGEKHKIKADKLVDLSANGITDVNSGKRAKTHILFR 536
Qy 444 GKQICNVACDPRDRVERHET-----LSQWRNRFSGA-----GF 477
Db 537 -----TPLLTPCTEHRVQTKVEYITKLNIRVDSKKITDGAASFTDLTNVQIRIGI 591
Qy 478 AAHIGS-NAPKQASMLLAL-----FNGGEGY-RVEESDG 510
Db 592 ELDNAGNVTKTKETKIIAKLGEQDNDVFGSGTTEIDGEGGYDRVHYSG 641

RESULT 10
SPYA_CALJA
ID SPTA_CALJA STANDARD; PRT; 414 AA.
AC P31029;
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SERINE-PYRUVATE AMINOTRANSFERASE, MITOCHONDRIAL PRECURSOR
DE (EC 2.6.1.51) (SPT) (ALANINE-GLYOXYLATE AMINOTRANSFERASE)
DE (EC 2.6.1.44) (AGT).
GN AGT OR AGT1.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix.
OX NCBI_TaxID=9483;
RN [1]
RX MEDLINE=92339467; PubMed=1339350;
RA Purdie P.E., Lumb M.J., Danpure C.J.;
RT "Molecular evolution of alanine/glyoxylate aminotransferase 1
in cellular targeting. Analysis of the marmoset and rabbit genes.";
RL Eur. J. Biochem. 207:757-766(1992).
CC -!- FUNCTION: DUAL METABOLIC ROLES OF GLUCONEOGENESIS (IN THE
MITOCHONDRIA) AND GLYOXYLATE DETOXIFICATION (IN THE PEROXISOMES).
CC -!- CATALYTIC ACTIVITY: L-SERINE + PYRUVATE = 3-HYDROXYPYRUVATE +
L-ALANINE.
CC -!- CATALYTIC ACTIVITY: L-ALANINE + GLYOXYLATE = PYRUVATE + GLYCINE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX AND PEROXISOMAL.
CC -!- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT
AMINOTRANSFERASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; M84414; AAA35397.1; -.
DR PIR; S24154; S24154.
DR InterPro; IPR000192; Aminotransf_class_v.
DR Pfam; PF00266; aminotran_5; 1.
DR PROSITE; PS00595; AA_TRANSFER_CLASS_5; 1.
KW TRANSFERASE; Aminotransferase; Pyridoxal phosphate; Peroxisome;
KW Mitochondrion; Transit peptide; Alternative initiation.
FT TRANSIT 1 23 MITOCHONDRION
FT CHAIN 24 414 SERINE-PYRUVATE AMINOTRANSFERASE,
FT MITOCHONDRIAL ISOFORM.
FT CHAIN 23 414 SERINE-PYRUVATE AMINOTRANSFERASE,
FT PEROXISOMAL ISOFORM.
FT INT_MET 23 23 FOR PEROXISOMAL ISOFORM.
FT BINDING 231 231 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 414 AA; 45054 MW; 604866DA42EEDDE1 CRC64;

Query Match 3.7%; Score 101; DB 1; Length 414;
Best Local Similarity 21.9%; Pred. No. 4.7;
Matches 70; Conservative 41; Mismatches 115; Indels 94; Gaps 15;

Qy 292 LRPGPPVFRITGTPPADN-----FDYLHEVGCKLA-HLAELHVEFEYR 337
Db 37 LKPLSIPTRLILG---PGSNLPPTMAAGGLQMLGHMHKETYQIMDEIKEGIQVFTFR 93
Qy 338 GFVANTLAD-----LDASMLE-LRPSEIESVAVNSVF-----ELHKLGRPG 378
Db 94 NPLTLVLSGSGHCALEALINVLPGSFLGVNGINGQRAADIGERLARGVHPMTKDPG 153
Qy 379 ---AIDKVLGVVNOIKPEIFTVVEQSNHNSPIFLDRFETSLHYSTLF-----DSLEGV 430
Db 154 GHYTLQVEVEGLAQHKPVLLFLTHGESSGVLQPLDGLGELCHRYKCLLLVDSVASLGA 213
Qy 431 P-----SGQDKVM-----SEVYLG-KQICNV 451
Db 214 PLYMDQGGIDILYSGSQKVLNAPPGTSLLSFSDTAKNKIYRKTKPSSFYLDVKYLANLW 273
Qy 452 ACDGPDPRVERHET---LSQWRNRFSGAFAAHIGSNFAKQASMLLALFNG---GEGYRVE 506
Db 274 GCDGQPMYHHTTPVSVLSYSLREGALLSLEOGL-ENSWRKHREAAAYLHGRLOALGLRLF 332
Qy 507 ESDGCLMLGWHTRPLIATSA 526
Db 333 VKDPALRL-----PTVTVA 347

RESULT 11
MA2B_FELCA
ID MA2B_FELCA STANDARD; PRT; 1007 AA.
AC O46432;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LYSOSOMAL ALPHA-MANNOSIDASE PRECURSOR (EC 3.2.1.24) (MANNOSIDASE,
ALPHA B) (LYSOSOMAL ACID ALPHA-MANNOSIDASE) (LAMAN).
GN MAN2B1 OR MANB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RX MEDLINE=98060744; PubMed=9396732;
RA Berg T., Tollersrud O.K., Walkley S.U., Siegel D., Nilssen O.;
RT "Purification of feline lysosomal alpha mannosidase, determination of
its cDNA sequence and identification of a mutation causing alpha-
mannosidosis in Persian cats.";
RL Biochem. J. 328:863-870(1997).
CC -!- FUNCTION: NECESSARY FOR THE CATABOLISM OF N-LINKED CARBOHYDRATES
RELEASED DURING GLYCOPROTEIN TURNOVER (BY SIMILARITY).
CC
```





```
QY 303 TGIGPPADP-----NFDYLHEVGCKLAHLAEAIHVEFEYRGV-----ANT 343
DQ 771 --LSPTEKIMKRRDTPEDIFVTF---IVHSSKLLKNE-GYLGFIPSPSFGTGVRYSNL 824
QY 344 LADLDASM---LRLPSEIESVAV--NSVFLHKLGRPGAKDKVL-----GVVN 388
DQ 825 RKELETKMCKLKLILYLPDPVFGAYVDCIILHK--RPPKSEDILVLYAPPKTKKISF 882
QY 389 QKQPEFTVVEQESNHNPS-----IFLDRF-----TESLHYSTLFDLSLEGVPS 432
DQ 883 EFKNDLF--IEYSKILNDPKCRIFPKSPDIYIILDKIKONCRESLYTLEDLTETIGILA 940
QY 433 GO-----DKVMSEVYLGKICNVVACDPRVERHETLSQWRNRFSGAFAAHIGSNAPK 488
DQ 941 SKYFESDKKENYLYLPYLEGVNY-----RYETKLKLN-----YVDFSKHK 981
QY 489 QASMLLALFNGE 501
DQ 982 NNEKLINLFMSPE 994

RESULT 13
DPO3_STA00 STANDARD; PRT; 1436 AA.
AC Q53665; Q57110; Q9F1J9;
DT 15-DEC-1998 (Rel. 37, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE DNA POLYMERASE III POLC-TYPE (EC 2.7.7.7) (POLIII).
GN POLC.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP8;
RX MEDLINE=96084951; PubMed=7489915;
RA Pacitti D.F., Barnes M.H., Li D., Brown N.C.;
RT "Characterization and overexpression of the gene encoding
RT Staphylococcus aureus DNA polymerase III.";
RT Gene 165:51-56(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Inoue R., Kaito C., Tanabe M., Kamura K., Akimitsu N., Sekimizu K.;
RT "Genetic identification of two distinct DNA polymerases, DnaE and
RT POLC, essential for chromosomal DNA replication in Staphylococcus
RT aureus.";
RN Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR REPLICATIVE DNA SYNTHESIS. THIS DNA
CC POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC N PYROPHOSPHATE + DNA(N).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. POLC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D86727; BAA13160.1; -
CC EMBL; Z48003; CAA88043.1; -
CC EMBL; AB053353; BAB20885.1; -
CC InterPro; IPR000520; Exonuclease.
CC InterPro; IPR003141; PHP_N.
CC Pfam; PF00929; Exonuclease; 1.
CC Pfam; PF02231; PHP_N; 1.
```

```
DR SMART; SM00479; EX0111; 1.
KW SMART; SM00481; POLIIIAAC; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; Hydrolase;
KW Nuclease; Exonuclease.
FT DOMAIN 420 584
FT CONFLICT 207 208 NE -> QO (IN REF. 1).
FT CONFLICT 952 952 T -> K (IN REF. 1).
FT CONFLICT 1030 1049 MISSING (IN REF. 1); CAA88043.1.
FT CONFLICT 1035 1035 R -> A (IN REF. 1).
FT CONFLICT 1147 1153 EFGTGFV -> NSDRIR (IN REF. 1).
FT CONFLICT 1260 1260 M -> I (IN REF. 1).
FT CONFLICT 1409 1409 S -> Y (IN REF. 1).
SQ SEQUENCE 1436 AA; 162459 MW; 2F70E034C0FBF723 CRC64;

Query Match 3.6%; Score 99.5; DB 1; Length 1436;
Best Local Similarity 20.9%; Pred. No. 37;
Matches 116; Conservative 84; Mismatches 206; Indels 149; Gaps 30;

QY 10 QDKKTMNEEDDG--NGMDLLAVLYGKYRSESMADYAKLEQLEVMMSNVQEDDLSQL 67
DQ 416 KDATVVFVDVETGLSNQYDKIIELAAVKVHNGEIID--KFER---FSNPHE-RLSET 467
QY 68 ATEVYHNPALYTWLDSMLTDLNPSSNAEYDLKAIPGDAILNOFAIDSSASSSSQGGG 127
DQ 468 IINLTHIT-----DDMLYDA-PEIEVLTETFEKEMWGDAL---FVAHNASFDM---G 511
QY 128 DTYTNKRL---KCSNGVET---TTATAESTRHVVLVDSOENGVRV--HALLACAEA 178
DQ 512 FIDTGYERLFGPSTNGVIDTLELSRTINTEYKHLNLAFLAKKYGVLTQHRAIYDTEA 571
QY 179 VOKENLTVAEALVKOIGFLAVSQIGAMRKVATYFAEALARRIYRLSPSQSPIDH---SLS 235
DQ 572 TAYIFIKMVQOM--KELGVLNHNKLNEDAYKRARPSSHVTLIVQOGLKNLFKIVS 630
QY 236 DTLQMHFVETCPYLKFAHFTANOALIEAF-----QGKRVHVVIDFSMSQ----- 279
DQ 631 ASLVKYFYRT-PRIP-----RSLDEYREGLLVGTACDEGELFTAVMQKDSQVEKIA 682
QY 280 -----GLQWALMALRPGPPVFLRTIGIGPPADPNFDYLVHEVGCKLAHLAEA--- 329
DQ 683 KYDFIEIQPALYQDLIDR---ELIRD-----ETLHEIYQRLIHAGDTAGIP 728
QY 330 -----IHVEFEYRGFVANTLA-----DLDSMLELRPSIEISVAVNSVFLHKLGRP 377
DQ 729 VIATGNAYLFEHGDGIARKILIASQPGNPLNRSTL---PEAHFRTDDEMLNEFH-FLGEE 784
QY 378 GAIDKVLGVNQIKPEITFTVVEQESNHNPSIFLDRFETSLHYSTLFDLSLEGVPSGDKV 437
DQ 785 KAHEIVKNTNELADRIERVV-----PIKDELYTPRM-----EGANEE 822
QY 438 MSEV-----YLGKQICNVVACDPRVERHETLSQWRNRFSGAFAAHIGSNAPKQA 490
DQ 823 IRELSYANARKLYGEDLPQIVI---DLEK-----ELKSIIGN-GFAVIYLIISORLVKK 872
QY 491 SMLLALFNGEGYRV 505
DQ 873 SL-----DDGYLV 880

RESULT 14
SPYA_FELCA STANDARD; PRT; 414 AA.
ID SPYA_FELCA
AC P41689;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SERINE--PYRUVATE AMINOTRANSFERASE, MITOCHONDRIAL PRECURSOR
DE (EC 2.6.1.51) (SPY) (ALANINE--GLYOXYLATE AMINOTRANSFERASE)
DE (EC 2.6.1.44) (AGT).
GN AGXT OR AGT1.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=94222101; PubMed=8168541;  
 RA Lumb M.J., Purdue P.E., Danpure C.J.;  
 RT "Molecular evolution of alanine/glyoxylate aminotransferase 1  
 intracellular targeting. Analysis of the feline gene.";  
 RL Eur. J. Biochem. 221:53-62(1994).  
 CC -!- FUNCTION: DUAL METABOLIC ROLES OF GLUCONEOGENESIS (IN THE  
 MITOCHONDRIA) AND GLYOXYLATE DETOXIFICATION (IN THE PEROXISOMES).  
 CC -!- CATALYTIC ACTIVITY: L-SERINE + PYRUVATE = 3-HYDROXYPYRUVATE +  
 L-ALANINE.  
 CC -!- CATALYTIC ACTIVITY: L-ALANINE + GLYOXYLATE = PYRUVATE + GLYCINE.  
 CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (>90%) AND PEROXISOMAL.  
 CC -!- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 AMINOTRANSFERASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X75923; CAA53527.1; -;  
 DR InterPro; IPR000192; Aminotransf\_class\_v.  
 DR Pfam; PF00266; aminotran\_5; 1.  
 DR PROSITE; PS00342; MICROBODIES\_CTER; 1.  
 DR PROSITE; PS00595; AA\_TRANSFER\_CLASS\_5; 1.  
 KW Transferase; Aminotransferase; Pyridoxal phosphate; Peroxisome;  
 KW Mitochondrion; Transit peptide; Alternative initiation.  
 FT TRANSIT 1 23 MITOCHONDRION (BY SIMILARITY).  
 FT CHAIN 24 414 SERINE--PYRUVATE AMINOTRANSFERASE,  
 FT CHAIN 24 414 MITOCHONDRIAL ISOFORM.  
 FT CHAIN 23 414 SERINE--PYRUVATE AMINOTRANSFERASE,  
 FT CHAIN 23 414 PEROXISOMAL ISOFORM.  
 FT INIT\_MET 23 23 FOR PEROXISOMAL ISOFORM.  
 FT BINDING 231 231 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT SITE 412 414 MICROBODY TARGETING SIGNAL (BY  
 SIMILARITY).  
 FT SEQUENCE 414 AA; 45507 MW; 0D1B01E0E9A199B3 CRC64;  
 Query Match 3.6%; Score 99; DB 1; Length 414;  
 Best Local Similarity 21.9%; Pred. No. 6.6; Mismatches 113; Indels 94; Gaps 15;  
 Matches 70; Conservative 43;  
 Qy 292 LRPGPPVFRLTGIGPP--AP-----DNFDYLHEVGCKLAHLAEAIHVEFEVGFVA 341  
 Db 37 LRPLSIPNLLGPGSNLAPRVLVAGGKQMGHMK---EMFQIMDDIKGQIVFQTK 93  
 Qy 342 NTL-----ADLDASMLE-LRPSETESVAVNSVF-----ELHKLGRPG 378  
 Db 94 NPLTLAISGSGHCALEALFNLEPGDPFLVGVNGVQRAADICERIGARVHPMWDPG 153  
 Qy 379 ---AIDKVLGVNQIKPEIFVVGQESHNSPIFLDRETSLSHYSTLF-----DSLEGV 430  
 Db 154 NHYTLQLEELAEALQHPVLLFLTOGESSGVLQPLDYGELCHRYNCLLLVDSVASLCGT 213  
 Qy 431 P-----SGQDKVMS-----EVYLG-KQICNVV 451  
 Db 214 PIYMQOQIDVLYSGSQVNSPPGTSLSFSKAKNKIYTRKTPVSYFLDMKLANIW 273  
 Qy 452 ACDGPDVRVERHET--LSQWRNRFSGAPFAAAHIGSNFAKQASMLLALFNG---GEGYRVE 506  
 Db 274 GCDGKPRIYHHTTPVVSLSLSRESLALAEQGL-ENSWRQREVNTAYLHGRLOGLGQLF 332  
 Qy 507 ESDGCLMLGWTRPLIATSA 526

Db 333 VKDPALRL-----PTVTVA 347  
 RESULT 15  
 PHBC\_RH1ET STANDARD; PRT; 636 AA.  
 AC Q52728;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE POLY-BETA-HYDROXYBUTYRATE POLYMERASE (EC 2.3.1.-) (POLY(3-  
 HYDROXYBUTYRATE) POLYMERASE) (PHB POLYMERASE) (PHA SYNTHASE)  
 DE DE (POLYHYDROXYALKANOATE) POLYMERASE (PHA POLYMERASE) (PHA SYNTHASE)  
 DE DE (POLYHYDROXYALKANOIC ACID SYNTHASE).  
 GN PHBC OR PHAC.  
 OS Rhizobium etli.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=29449;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CE 3;  
 RX MEDLINE=96198168; PubMed=8626293;  
 RA Cevallos M.A., Encarnacion S., Leija A., Mora Y., Mora J.;  
 RT "Genetic and physiological characterization of a Rhizobium etli mutant  
 strain unable to synthesize poly-beta-hydroxybutyrate.";  
 RL J. Bacteriol. 178:1646-1654(1996).  
 CC -!- FUNCTION: POLYMERIZES D(-)-3-HYDROXYBUTYRYL-COA TO CREATE PHB  
 WHICH CONSISTS OF THOUSANDS OF HYDROXYBUTYRATE MOLECULES LINKED  
 END TO END. PHB SERVES AS AN INTRACELLULAR ENERGY RESERVE MATERIAL  
 WHEN CELLS GROW UNDER CONDITIONS OF NUTRIENT LIMITATION.  
 CC -!- PATHWAY: THIRD STEP IN POLY-BETA-HYDROXYBUTYRATE BIOSYNTHESIS.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE PHA/PHB SYNTHASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U30612; AAB05020.1; -;  
 DR InterPro; IPR000073; Abhydrolase.  
 DR Pfam; PF00561; abhydrolase\_1.  
 KW PHB biosynthesis; Transferase; Acyltransferase.  
 FT ACT\_SITE 373 373 POTENTIAL.  
 FT SEQUENCE 636 AA; 71935 MW; DD370C10ECC89D12 CRC64;  
 Query Match 3.6%; Score 99; DB 1; Length 636;  
 Best Local Similarity 21.6%; Pred. No. 12;  
 Matches 80; Conservative 50; Mismatches 138; Indels 102; Gaps 18;  
 Qy 15 MMNNEEDGNGMDLAVLGKYKRSSMADVAQKLEQVYMMNVQDDLSQL-----AT 69  
 Db 233 MKMLAEDIAAGKGL-----RLRQTDMTKFAVGRDMALTPGKVIQNDICQIIQYEAST 286  
 Qy 70 ETVIYNPAELY-TWLDNM-LTDLNPSS-----NAYEDLKAI 104  
 Db 287 ETVLKRLPILICPPWINKFYILDNLNPKQSFIRKWCVDQGTVFISWNPDPGRHAEKDMAAY 346  
 Qy 105 PGDAILNQFAIDSASSNSQGGDPTTNNKRLKCSNGVWETTTATAESTRHVVLDVSOEN 164  
 Db 347 AREGI--DFALETTEKA-----TGEKVNAGVCGVGTLLATLALH-----AREK 390  
 Qy 165 GVRUVHALLACAE-----VOKENLTVAEALVKQIGFL-----AVSQIGAMR 206  
 Db 391 NKRIKTATLFTQVDFTHAGDLKVFVDEEQLAALAEHMQAAGYLDGSKMSMAFNMLRASE 450  
 Qy 207 KVATYFAEALARRIYRLSPSQSPIDHSL-----SDFLOM-----HFYETCPYLKFAHTAN 257

